

Parfit on theories of personal identity

PHIL 20208
Jeff Speaks

October 24, 2006

1	Parfit on reductionism vs. non-reductionism	1
2	Two versions of reductionism	2
2.1	The physical criterion	2
2.2	The psychological criterion	2
2.2.1	The transitivity objection (p. 206)	3
2.2.2	Memory presupposes rather than explains personal identity (pp.219-222)	3
2.2.3	Experiences vs. the subject of experiences (p. 222-224)	4
2.2.4	Williams' argument (pp. 229-243)	4

1 Parfit on reductionism vs. non-reductionism

Parfit draws a distinction between two different kinds of views about a certain thing. According to a *non-reductionist* view of something, the existence of that kind of thing is a 'further fact', which goes beyond the existence of other facts, not about the existence of that kind of thing. According to a *reductionist* view of something, the existence of that thing consists in the existence of certain other things.

Parfit gives two plausible examples of things about which we naturally have a reductionist view: nations and clubs.

Two possible examples of things about which we naturally have a non-reductionist view: elementary particles and, if you believe in the possibility of zombies and certain kinds of spectrum inversion, qualia.

Given this distinction, Parfit then raises an important question: should we be reductionists or non-reductionists about *persons*?

An example of a non-reductionist view of persons is that we are immaterial souls of the kind envisaged by the dualist. Parfit thinks that this kind of view is wrong, and that we should be reductionists about persons: the existence of persons is not a further fact, beyond the psychological and physical facts. He gives two kinds of arguments against this sort of dualist view:

- We do not directly observe persons as things above and beyond the physical and psychological facts which we do observe (pp. 223-224); all we are aware of is a certain kind of psychological connectedness. (In this, Parfit agrees with Hume.) So, if we believe in non-reductionist persons, we should require some evidence for this. And there could have been evidence, of at least the following two kinds: (1) we could have had very strong evidence for reincarnation, as in the example of the Japanese woman and the Celtic hunter (p. 227); (2) we could have had evidence that changes in a person's brain change their personality in an 'all or nothing' way (p. 228). But we do not find either such kind of evidence.
- It is hard to imagine cases in which we hold everything in the world fixed, but swap the identities of persons. This indicates that the existence of persons is not a 'further fact' which goes beyond the facts which we hold fixed.

We will return to this kind of non-reductionist, dualist view next week. But for now, let's grant that Parfit has made some challenging objections to the dualist position, and see where he goes with the view that reductionism about persons is the correct view.

2 Two versions of reductionism

In giving a reductionist view of personal identity, one specifies the facts which the existence of persons are 'nothing over and above.' The two natural candidates here are: physical facts about spatiotemporal continuity, and psychological facts.

2.1 *The physical criterion*

Parfit describes what he calls 'the standard view' of the existence of ordinary material objects, like billiard balls (p. 203). Applied to persons, this is the view that the existence of persons over time consists in

"the physical continuity, over time, of my brain and body. I shall continue to exist if and only if this particular brain and body continue to exist and to be the brain and body of a living person." (204)

2.2 *The psychological criterion*

In most of our discussion of personal identity so far, we have been focusing on various versions of the psychological view. We have encountered, from Reid and Williams, a number of serious objections to this view. Parfit provides responses to a number of these objections in trying to find the most plausible version of the psychological criterion.

2.2.1 *The transitivity objection (p. 206)*

Parfit takes Locke's view as the basis for his own. He begins discussion of it by considering an objection based on the following objection to Locke's view from Reid:

“There is another consequence of this doctrine, which follows no less necessarily, though Mr. Locke probably did not see it. It is, *that a man may be, and at the same time not be, the person that did a particular action.*

Suppose a brave officer to have been flogged when a boy at school for robbing an orchard, to have taken a standard from the enemy in his first campaign, and to have been made a general in advanced life; suppose, also, which must be admitted to be possible, that, when he took the standard, he was conscious of his having been flogged at school, and that, when made a general, he was conscious of his taking the standard, but had absolutely lost the consciousness of his flogging.

These things being supposed, it follows, from Mr. Lockes doctrine, that he who was flogged at school is the same person who took the standard, and that he who took the standard is the same person who was made a general. Whence it follows, if there be any truth in logic, that the general is the same person with him who was flogged at school. But the generals consciousness does not reach so far back as his flogging; therefore, according to Mr. Lockes doctrine, he is not the person who was flogged. Therefore the general is, and at the same time is not, the same person with him who was flogged at school.”

The problem, in general, is that identity is transitive, whereas direct memory connections are not. Parfit's solution of this problem in terms of the distinction between direct psychological connectedness and psychological continuity. Though the former is not transitive, the latter is; so Reid's objection to analyses of personal identity in terms of the latter does not hold.

2.2.2 *Memory presupposes rather than explains personal identity (pp.219-222)*

Later, Parfit considers another objection to his version of the psychological view, which can again be found in Reid's objections to Locke:

“It may here be observed (though the observation would have been unnecessary, if some great philosophers had not contradicted it), that it is not my remembering any action of mine that makes me to be the person who did it. This remembrance makes me to know assuredly that I did it; but I might have done it, though I did not remember it. That relation to me, which is expressed by saying that I did it, would be the same, though I had not the least remembrance of it. To say that my remembering that I did such a thing, or, as some choose to express it, my being conscious that I did it, makes me to have done it, appears to me as great an absurdity as it would be to say, that my belief that the world was created made it to be created.”

Parfit's response in terms of the notion of a 'quasi-memory.'

2.2.3 *Experiences vs. the subject of experiences (p. 222-224)*

A third objection which Parfit considers is based on Reid's objection that the psychological theory ignores the fact that the existence of a person is not based on continuity of experiences, but on continuity of the subject of those experiences, which is something above and beyond those experiences themselves:

"My personal identity, therefore, implies the continued existence of that indivisible thing which I call myself. Whatever this self may be, it is something which thinks, and deliberates, and resolves, and acts, and suffers. I am not thought, I am not action, I am not feeling; I am something that thinks, and acts, and suffers."

Parfit's objections to the idea that we are aware of a self, in this sense.

2.2.4 *Williams' argument (pp. 229-243)*

The most important objection to the psychological view for the purposes of understanding Parfit's position is the objection which Williams gives in 'The self and the future.'

Parfit's reworking of Williams' argument in terms of the 'psychological spectrum', which is a continuum of very many cases which stretch, on the one hand, from replacement of one or two of my memories with those of Napoleon, to a case on the other end of the spectrum, on which all of my memories are replaced with those of Napoleon.

There seem to be three things we can say about this range of cases:

- We should say the same thing about these cases as we are inclined to say about the structurally similar 'paradox of the heap', or 'sorites paradox': sometimes the question about whether a person which results from some change to me *is* me is an empty question, which can only be answered by stipulation.
- We should say that there is a sharp dividing line somewhere in the psychological spectrum, so that if, say, 53.4% or more of my memories are changed, then the resulting person is not me, but that if less are changed, that person is me.
- We should say that in every case, the resulting person is me.

As Parfit says, the force of Williams' argument resides largely in the fact that (3) seems like the most plausible answer here. We find the view that personal identity can sometimes be indeterminate or arbitrary hard to accept, and it also seems hard to accept that there is some principled dividing line in this spectrum of cases. So it seems plausible to conclude, with Williams, that personal identity is consistent with any amount of psychological change. This would show that the psychological criterion of personal identity is incorrect.

And, if we are committed to reductionism, this would seem to show that the physical criterion must be correct.

Though Parfit sees the force of this argument, he does not agree with Williams' conclusion. He shows that a similar argument can be run against the physical criterion of personal identity using the continuum of cases that he calls the 'physical spectrum,' which go from, on the one hand, a replacement of one or two of the cells in my brain and body with exact duplicates to, on the other hand, a replacement of all the cells in my brain and body with exact duplicates. We seem to face a similar range of choices as above:

- We should say the same thing about these cases as we are inclined to say about the structurally similar 'paradox of the heap', or 'sorites paradox': sometimes the question about whether a person which results from some change to me *is* me is an empty question, which can only be answered by stipulation.
- We should say that there is a sharp dividing line somewhere in the physical spectrum, so that if, say, 53.4% or more of my cells are changed, then the resulting person is not me, but that if less are changed, that person is me.
- We should say that in every case, the resulting person is me.

As above, the most appealing option seems to be (3). But if this is true, then the physical criterion must be false, since then personal identity is consistent with complete physical discontinuity.

We seem at this point to be in a tough spot. We seem to have only two options: the physical and psychological views. But now we seem to have convincing arguments against both, since neither seems to provide a necessary condition for personal identity.

We could accept the conclusions of both arguments, and adopt a *mixed* view. On this view, either physical continuity *or* psychological continuity would be sufficient for personal identity, and it would only be necessary to have one or the other.

Why the *combined spectrum* seems to rule this view out.

This seems to leave us with only two choices: there is a sharp dividing line somewhere in the combined spectrum, or sometimes there is no fact of the matter about personal identity, and the question 'Is x the same person as y ?' is an empty question.

Can you think of any criterion of personal identity which can survive these three 'spectrum' cases?